JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2011.02339.x

Determinants of distribution and prevalence of avian malaria in blue tit populations across Europe: separating host and parasite effects

E. SZÖLLŐSI*, M. CICHOŃ†, M. EENS‡, D. HASSELQUIST§, B. KEMPENAERS¶, S. MERINO**, J.-Å. NILSSON§, B. ROSIVALL*, S. RYTKÖNEN††, J. TÖRÖK*, M. J. WOOD‡‡[1] & L. Z. GARAMSZEGI§§

*Behavioural Ecology Group, Department of Systematic Zoology and Ecology, Eötvös Loránd University, Budapest, Hungary
†Institute of Environmental Sciences, Jagiellonian University, Kraków, Poland
‡Ethology Research Group, Department of Biology, University of Antwerp, Wilrijk, Antwerp, Belgium
§Department of Biology, Lund University, Lund, Sweden
¶Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany
**Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Científicas, Madrid, Spain
††Department of Biology, University of Oulu, Oulu, Finland
‡‡Edward Grey Institute, Department of Zoology, University of Oxford, Oxford, UK
§§Department of Evolutionary Ecology, Estación Biológica de Doñana-CSIC, Seville, Spain

Keywords:

blood parasite;
Cyanistes caeruleus;
geographical distribution;
Haemoproteus;
host specificity;
Parahaemoproteus;
Parus;
Plasmodium;
specialization.

Abstract

Although avian malarial parasites are globally distributed, the factors that affect the geographical distribution and local prevalence of different parasite lineages across host populations or species are still poorly understood. Based on the intense screening of avian malarial parasites in nine European blue tit populations, we studied whether distribution ranges as well as local adaptation, host specialization and phylogenetic relationships can determine the observed prevalences within populations. We found that prevalence differed consistently between parasite lineages and host populations, indicating that the transmission success of parasites is lineage specific but is partly shaped by locality-specific effects. We also found that the lineage-specific estimate of prevalence was related to the distribution range of parasites: lineages found in more host populations were generally more prevalent within these populations. Additionally, parasites with high prevalence that were also widely distributed among blue tit populations were also found to infect more host species. These findings suggest that parasites reaching high local prevalence can also realize wide distribution at a global scale that can have further consequences for host specialization. Although phylogenetic relationships among parasites did not predict prevalence, we detected a close match between a tree based on the geographic distance of the host populations and the parasite phylogenetic tree, implying that neighbouring host populations shared a related parasite fauna.

Correspondence: Eszter Szöllősi, Behavioural Ecology Group, Department of Systematic Zoology and Ecology, Eötvös Loránd University, Pázmány P. sétány 1/C, H-1117 Budapest, Hungary.
Tel.: +36 1 381 2193; fax: +36 1 381 2194;
e-mail: sz_eszter@ludens.elte.hu
[1]Present address: Department of Natural and Social Sciences, University of Gloucestershire, Francis Close Hall, Cheltenham GL50 4AZ, UK.

Introduction

A central issue in understanding host–parasite dynamics is to determine what factors shape prevalence (i.e. the parasites' ability to infect and transmit between hosts; Dybdahl & Storfer, 2003). Although there is a continual evolutionary arms race between host and parasite, both

parties should optimize the resource allocation in a way that maximizes their fitness. Hence, detected prevalence should reflect the outcome of such conflicts, and for understanding the parasite's distribution across host individuals, one should take into account several factors that may influence the success of parasite transmission across hosts. For example, host specificity of parasites (Hellgren *et al.*, 2009), the way of parasite transmission (Ewald, 1983) and the duration of a given host–parasite coevolutionary relationship (Ewald, 1983; Dybdahl & Storfer, 2003; Poulin & Mouillot, 2004; Garamszegi, 2011) as well as constrains due to the phylogenetic relationships of hosts and parasites (Beadell *et al.*, 2009; Hellgren *et al.*, 2009) may each influence prevalence.

A long coevolutionary relationship with a host species may result in a better host exploitation strategy by the parasite, which may manifest in higher parasite prevalence in this specific host (Poulin & Mouillot, 2004). However, closer adaptation to the host defence mechanisms may also result in a cost to the parasite, namely the loss of genetic variation associated with the ability to use alternative hosts. Furthermore, the development of evasion mechanisms against additional host species is also expected to be costly to the parasite (Combes, 1997). Therefore, a generalist parasite that infects several distantly related host species may be less able to adapt to all its host species and thus will show lower prevalence (and virulence) than specialist parasites (trade-off hypothesis: Ewald, 1983; Poulin, 1998; Garamszegi, 2006). In support of this hypothesis, greater taxonomic distance between the utilized host species has been found to associate with lower infection success in Cestodes and Nematodes (Poulin & Mouillot, 2004).

The resource breadth hypothesis (Brown,1984), on the other hand, emphasizes that the attribute determining the local abundance of a species also affects its distribution. Accordingly, and contrary to the trade-off hypothesis, a positive correlation between the number of potential host species and parasite prevalence can be expected, as has been shown by two previous studies. Siphonaptera fleas were more abundant and avian malarial parasites achieved higher prevalence if they exploited numerous unrelated hosts as compared with more specialist parasites infecting fewer hosts (Krasnov *et al.*, 2004; Hellgren *et al.*, 2009). One explanation could be that a broader host range may increase the transmission success of generalist parasites in habitats with a diverse array of host species, because the possibility of encountering at least one suitable host is higher than for parasites exploiting a single host. This hypothesis also applies to vectorborne parasites, such as malarial parasites, because vectors of these parasites also rely on a range of host species (Malmqvist *et al.*, 2004), thus allowing for a parasite to be transmitted to different host species. Note that patterns supporting the resource breadth hypothesis not only may result from host specialization (i.e. host ranges) but can also be caused by differences in host geographic ranges.

The aforementioned hypotheses rely on the inherent but so far untested assumption that parasite prevalence is a species- or lineage-specific trait. However, locally detected prevalence can be shaped by various factors such as vector and host density or climatic conditions (Wood *et al.*, 2007; Merino *et al.*, 2008). The degree by which prevalence is mediated by parasite-specific and local effects is currently unknown. Until these factors are not separated, it should not be taken granted that the same parasite will necessarily have similar prevalence in different hosts and/or locations.

Phylogenetic constraints may also influence parasite prevalence. If local prevalence is an innate characteristic of parasite lineages or species due to a complex genetic machinery that determines the interaction between hosts, parasites and vectors, it is likely that closely related parasites will realize similar prevalence as a result of the similarity in their transmission and reproductive mechanisms. On the contrary, if prevalence is a more flexible trait on an evolutionary time scale, it can be predicted to vary independently of phylogeny.

Avian malarial parasites (order Haemosporidia) are pathogens of both domestic and wild birds, and as such, they have been extensively studied (Valkiūnas, 2005). Recent molecular studies suggested that thousands of lineages may exist, many of which can infect multiple host species (Ricklefs & Fallon, 2002; Waldenström *et al.*, 2002; Martinsen *et al.*, 2006; Hellgren *et al.*, 2007). A routine screening of parasites based on molecular tools offers new insights into the life-history evolution of avian malaria. For example, previous studies of host specificity and distribution of avian malaria sampled a range of host species and controlled for the phylogenetic relationships of both hosts and parasites (Ricklefs & Fallon, 2002; Beadell *et al.*, 2004; Ricklefs *et al.*, 2004; Hellgren *et al.*, 2009). However, there is little information concerning the consistency of the composition of the avian malarial parasite fauna among the populations of the same host species sampled at different localities (but see Kimura *et al.*, 2006), despite the potential for population differences to exist (Bonneaud *et al.*, 2006; Bowen *et al.*, 2006; Merino *et al.*, 2008). Similarly, the predictions of the trade-off and resource breadth hypotheses have never been tested on different populations of the same host species. It is therefore of considerable importance to examine the distribution and prevalence of parasite species among populations of the same host species. Moreover, the sampling of the same lineages in different locations allows estimating the degree by which prevalence is shaped by lineage-specific factors.

In this study, we aimed to test the relative importance of parasite phylogeny, lineage identity and geographical constraints in determining parasite distribution and prevalence in different populations of the same nonmigratory host species, the blue tit (*Cyanistes caeruleus*). First,

we examined whether lineage-specific or locality-specific attributes or both determine detected prevalences. If parasite prevalence is a lineage-specific trait, we predict that the prevalence of the lineages will show consistent variation across locations. If prevalence is principally determined by local environmental factors (and not by the parasite's characteristics), we predict that prevalence is more consistent within localities than across lineages.

Second, we tested for the relationship between mean prevalence across populations and geographic distribution (in terms of the number of populations infected). According to the resource breadth hypothesis, we predicted that those parasites would be more widely distributed between host populations that are also more prevalent within these populations. This prediction can also be extended to the interspecific level, and thus, we can test for a positive relationship between prevalence and the number of host species infected by the same parasite lineage. On the other hand, if the number of populations in which a given lineage occurs reflects host specificity, the trade-off hypothesis predicts a negative relationship between parasite prevalence and geographic distribution. It has previously been shown that genetic distance between blue tit populations increased with geographic distance and populations only a few kilometres away from each other were well differentiated (Verheyen *et al.*, 1995). Therefore, we can also assume that the blue tit populations sampled in this study are also genetically different, presumably showing different host–parasite coevolutionary patterns. As a consequence, we can regard the number of blue tit populations infected by a particular parasite as an estimate of host specificity, as wider distribution of parasites might reflect their ability to infect genetically distinct hosts. Our predictions hold also for the interspecific level, if the constraint of host specificity is expressed in the number of host species infected. Hence, we can predict that host range in terms of the number of host species that a parasite infects would negatively correlate with its prevalence.

Third, we also investigated phylogenetic constraints that affect prevalence and distribution range. If prevalence is similar in closely related parasite lineages, we should observe that prevalence is structured phylogenetically. Moreover, if the global distribution of parasites is shaped by the phylogenetic relationship between hosts and parasites, we predict that the evolution of parasite lineages will not vary independently of the spatial distribution of hosts.

Methods

Sample collection and parasite detection

Blood samples (10–20 μL) were collected from a total of 476 adult blue tits (*C. caeruleus*) from nine nest box breeding populations across Europe (Table 1, Fig. 1). The blue tit is a hole nesting, nonmigratory passerine, although it is a partial migrant in its northern distribution range (Nilsson *et al.*, 2008). Start of egg laying depends on latitude and altitude (Fargallo, 2004) but eggs are usually laid from late March until mid-May and blue tits may re-nest after fledging of their first brood. Clutch completion and incubation take approximately a month; thus, nestlings hatch from late April to mid-June. To control for potential seasonal and annual changes in the presence of different species of avian malarial parasites in the birds' blood, each host population was sampled in May–June 2005, when both parasite genera may be present in the blood (Valkiūnas, 2005). We also controlled for a possible sex-specific sensitivity of the host species (McCurdy *et al.*,1998) to malarial infection by sampling approximately the same number of adult males and females in each population (Table 1). Adult birds were sampled during the nestling feeding stage (i.e. when nestlings from the first clutch were 4–14 days old), thus providing a similar hormonal milieu for parasites in each host population (which otherwise may have biased our estimate of parasite prevalence, e.g. Escobedo *et al.*, 2005).

Blood samples were stored in absolute ethanol or in Queen's lysis buffer (Seutin *et al.*, 1991) and kept at −20 °C or at room temperature until laboratory analysis. After DNA extraction, the concentration of genomic DNA was adjusted to 50 ng μL^{-1}. Polymerase chain reactions (PCRs) were performed to amplify a part of the cyt*b* gene on the *mt*DNA of the parasites using the protocol described by Waldenström *et al.*, 2004. In all PCRs, both negative (ddH_2O) and positive controls (samples from birds which were previously confirmed to be infected) were included among the samples to control for possible contaminations and failures during PCRs, respectively. To ensure that none of the samples went through degradation between sample collection and analysis, all negative samples were checked for DNA quality by amplifying the CHD (chromo-helicase-DNA-binding) genes of the host DNA (Griffiths *et al.*, 1998).

All samples with positive amplification were sequenced directly using the BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA), and the products from the sequencing reactions were run on an ABI PRISM® 3100 Genetic Analyser (Applied Biosystems). Sequences were edited and aligned using the program BioEdit (Hall, 1999) and identified to genus level (and classified them to be *Parahaemoproteus* or *Plasmodium* based on the most recent phylogenetic study by Martinsen *et al.*, 2008) by comparing sequence data with those of previously identified parasites. Parasites with sequences differing by one nucleotide substitution were considered to represent evolutionary independent lineages (Bensch *et al.*, 2004).

Calculating lineage-specific traits of parasites: prevalence and geographic distribution across hosts

After the identification of parasites, we calculated parasite prevalence at each sampling location for each lineage

Table 1 Prevalence (%) of malaria lineages (the percentage of infected individuals in relation to all screened individuals in a population) in different blue tit populations. Location names refer to the closest settlement near the sampling sites. 'Pa' refers to lineages belonging to *Parahaemoproteus;* 'Pl' refers to *Plasmodium* in lineage names. GenBank accession numbers are indicated under the lineage names.

	Prevalence													
Location	Pa-PARUS1 AF254977	Pa-PARCAE1 HQ537480	Pa-BLUTI1 DQ991077	Pa-ROFI1 DQ060769	Pl-COLL1 AY831747	Pl-GRW11 AY831748	Pl-GRW6 DQ368381	Pl-SGS1 AF495571	Pl-TURDUS1 HQ537478	Pl-BT7 AY393793	Pl-SW2 AF495572	Pl-ACAGR1 HQ537479	Pl-BLUTI1 DQ991068	Sample size (males/ females)
Vienna, Austria (48°13′N, 16°20′E)	19.1				2.4	9.5	4.8	47.6						42 (21/21)
Antwerp, Belgium (51°09′N, 4°24′E)	2. 7							38.7						75 (36/39)
Oulu, Finland (65°03′N, 25°31′E)	10.4	3.5							31.0					29 (12/17)
Hemse, Sweden (57°10′N, 18°20′E)	22.0								48.8	17.1				41 (22/19)
Pilisszentlászló, Hungary (47°43′N, 19°01′E)						12.2		38.8						49 (25/24)
Revinge, Sweden (55°42′N, 13°28′E)	60.0			2.4					16.7	7.1	2.4			42 (21/21)
Krakow, Poland (50°06′N, 20°25′E)	28.6					3.2		31.8	3.2	1.6		1.6		63 (31/32)
San Ildefonso, Spain (40°53′N, 4°01′W)	100.0													40 (20/20)
Wytham, UK (51°47′N, 1°20′W)			1.1			2.1		10.5	9.5	5.3	1.1		1.1	95 (48/47)

Fig. 1 Geographical distribution of sampling sites. (The map is used with permission of Cartographic Research Lab at the University of Alabama.)

that is present at those locations as the number of infected birds/screened birds. Prevalence was arcsine-square-root-transformed for the statistical analyses to satisfy the requirement of normality. For lineages that were detected in more than one host population, multiple data on prevalence were available that allowed us to test for lineage-specific effects. Similarly, in all except the Spanish population, more than one parasite lineage was present, which permitted the assessment of population-specific effects. These effects were estimated in a two-way ANOVA model that included both lineage and locality as main factors. For this model, we only used lineages that were prevalent in at least two host populations. After identifying the determinants of malaria prevalence at the sampling level, from the same model, we calculated lineage-specific estimates of prevalence that were independent of host population effects in the form of least square (LS) means. LS means of prevalence were therefore obtained for lineages that were present in more than one host population. For lineages detected in a single locality, we used prevalence information from that locality without any correction. These combined estimates were utilized in a set of phylogenetic analyses to test for the determinants of malaria prevalence at the across lineage level (for similar approaches see Lucas *et al.*, 2004; Garamszegi, 2006). However, when we simply used the mean prevalence for each lineage across sampling locations, we obtained results that were qualitatively identical to those reported below.

Note that prevalence was calculated by considering lineages that were present in the given population. We did not include zero prevalence for two reasons. First, the meaning of parasite absence from a location is not obvious. The absence of a parasite from a location can indicate that the population was exposed to the parasite but the parasite could not spread into that location (meaningful zero) or that the population was not yet exposed to the parasite (meaningless zero). With the current data, it is impossible to distinguish between these two possibilities. Second, entering zero prevalence in all absence case for each lineage on each location would cause bias in our analyses. Specialist or narrowly distributed parasites that are thus absent in many locations would necessarily receive zero prevalence in those locations, which would drive a false-negative correlation between prevalence and host range. We found a strong correlation between minimum (with zeros) and maximum (without zeros) estimates of prevalence across lineages ($r = 0.790$, $n = 13$, $P = 0.001$), and thus we infer that any error that is caused by the imprecision of our estimate due to the exclusion of meaningful absence cases should be of minor magnitude. Given the high sensitivity of the PCR screening (Waldenström *et al.*, 2004) and the large sample size we obtained in each location (see Table 1), we also infer that if we did not detect a parasite in a location, it truly means that the parasite is absent on that location.

Phylogenetic analyses

Reconstruction of the phylogenetic history

We used the nucleotide sequences to generate phylogenetic trees based on Bayesian sampling as implemented in the software MRBAYES 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). We used *Plasmodium falciparum* as outgroup and recovered a large set of trees, in which the frequency of occurrence of particular trees in the sample was proportional to their data fit. This set was sampled from a Markov chain implementing a general time reverse model of evolution with gamma correction for heterogeneity among sites, as the GTR + Gamma model is widely used to model sequence evolution in malaria (e.g. Beadell *et al.*, 2004; Yotoko & Elisei, 2006; Hellgren *et al.*, 2009). As all Bayesian analyses, the Bayesian reconstruction of phylogenies requires the definition of priors, which define the distribution of parameters, in which estimated parameters are expected to occur. The chain used uniform prior probabilities on trees and the parameters of the model of sequence evolution and an exponential prior on branch length. We allowed the chain to reach convergence and then sampled 100 trees at intervals of 1000 trees (1 million iterations) after a stationary point (burn-in) that was identified based on the (i) plots of log-likelihoods over time, (ii) similarity in topologies, branch support (posterior probabilities, Pp) and log-likelihoods between trees from each replicate and (iii) the average standard deviation of split frequencies between runs.

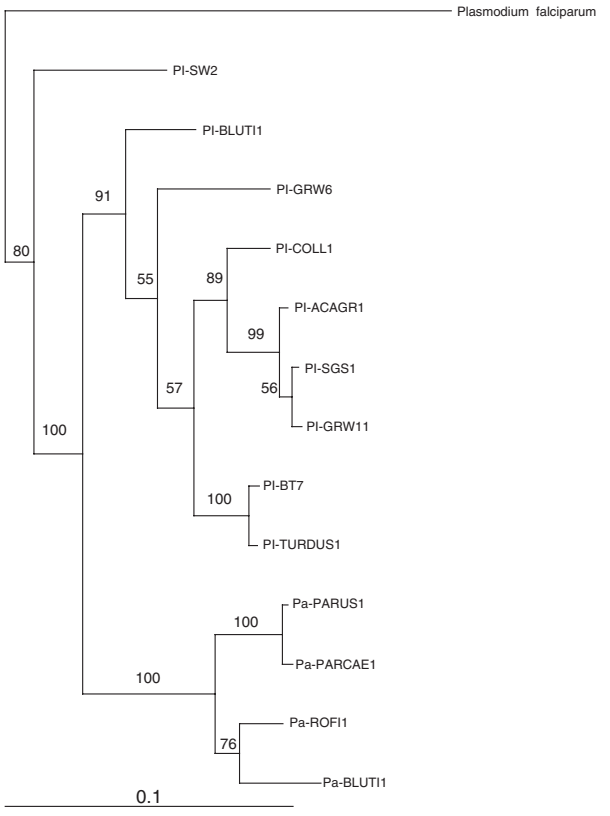

Fig. 2 The consensus phylogenetic tree of avian malarial parasites of blue tits originating from nine European populations. The phylogenetic reconstruction was based on the Bayesian analysis of cytochrome *b* sequences, in which 100 trees were sampled at intervals of 1000 trees from a converged Markov chain. Numbers at the nodes indicate the Bayesian posterior probabilities of each partition or clade in the tree, which are the proportion of trees in the sample that have the particular node. Branch lengths reflect the expected nucleotide substitutions per site. [Correction added after online publication 7 July 2011: Labelling of Fig. 2 corrected.]

The phylogeny and branch lengths were estimated from the majority-rule consensus of the pooled post-burn-in trees from the two replicates (Fig. 2).

Modelling the evolution of prevalence

We used a Bayesian modelling of continuous character state evolution to estimate the phylogenetic constraints of prevalence (i.e. to test whether closely related lineages have similar prevalence in their host populations) and to identify the evolutionary role of parasite distribution in mediating prevalence. For the assessment of phylogenetic constraints and the role of correlated trait evolution, we used the Bayesian framework available in BayesTraits (Pagel *et al.*, 2004). The effect of phylogeny can be approximated through the estimation of the phylogenetic scaling parameter lambda (λ) that varies between 0 (phylogenetic independence) and 1 (lineages' traits covary in direct proportion to their shared evolutionary history) (Freckleton *et al.*, 2002). The effect of parasite distribution was monitored in the form of the phylogenetic correlation between parasite prevalence and the number of host populations infected (ranging from one to seven).

To test our predictions in relation to host range at the interspecific context, we used the number of host species. The number of host species that a parasite infects was extracted from the database MalAvi (Bensch *et al.*, 2009) from which only infection data for *Passeriform* birds were used, as nonpasserine hosts are disproportionately under-represented in the database.

The Bayesian modelling relied on the full Bayesian sample of phylogenetic trees that allowed us to control for any phylogenetic uncertainty. These trees were fitted to the lineage-specific data on prevalence and geographical distribution under the assumption of a 'random walk' evolutionary scenario (Model A). We allowed a Markov chain to run for 5 million cycles after convergence, which was assessed by comparing results across five runs and plotting time-series graphs (not shown). Each iteration provided a phylogenetic model, which can be specified by trait values (alpha) and their variances corresponding to the root of the phylogeny. These parameters were sampled every 100 generations after a burn-in of 50 000 iterations (log-likelihood, $r = 0.019$; $n = 50\,000$). This sample of evolutionary models was used to estimate the posterior distributions of the phylogenetic scaling factor (λ) and the phylogenetic correlation between parasite prevalence and geographical distribution.

The transition rate parameters of continuous-time Markov models of trait evolution were conditioned by adjusting the 'ratedev' parameter to a value that provides an acceptance rate of newly proposed states of the rate parameters between 20% and 40%. To assess the robustness of model convergence, we ran at least five independent Markov chains of 5.5 million observations, which all converged to the same direction providing very similar posterior distributions of rate parameters and ancestral state estimations. Chains were sampled at each 100th iteration, and burn-in was set to 50 000 resulting in a sample of 50 000 observations.

To estimate the importance of the phylogenetic relatedness of lineages via the phylogenetic scaling factor, we compared posterior distributions and model likelihoods from the set of models that constrained λ to be zero or one with posterior distributions from the set that allowed λ to be estimated. Similarly, we estimated the posterior distribution of the phylogenetic correlations investigated (e.g. between prevalence and geographic distribution and between geographic distribution and the number of host species) by checking model likelihoods and posterior distributions when the correlation was forced to be zero and when it was allowed to be estimated. For these comparisons of different model parameter settings, we focused on the harmonic mean

(H_{mean}) of the likelihoods estimated across a large number of iterations and at different model settings. H_{mean} approximates the marginal likelihood of a model, which is the integral of the model likelihoods over all values of the model parameters and over all possible trees. H_{means} can be used to calculate the Bayes factor (BF) statistic in a form of 2(log[H_{mean} (better model)] − log[H_{mean} (worse model)]), which can then be used for hypothesis testing. According to the convention, BF values greater than two can be interpreted as 'positive' evidence for the model with the higher H_{mean} explaining the data better than the model with lower H_{mean}, whereas values above five indicate 'strong' evidence (Pagel & Meade, 2006).

Spatially constrained evolution: coevolution with the host
To study the historical association between parasite lineages and host populations, the putative history of host–parasite associations needs to be reconstructed by comparing the phylogenetic relatedness of parasites with the phylogenetic relatedness of their hosts in the light of infection patterns. Because phylogenetic information was unavailable for the studied host populations, we used the geographic distance matrix of localities to reconstruct the pattern of relatedness between host populations by applying tree-clustering methods. We thus assumed that geographical distances between host populations reflect true genetic distances, which is generally the case (Verheyen *et al.*, 1995; Nagai *et al.*, 2007; Lindsay *et al.*, 2008). From the perspective of the parasites, we used the consensus tree from the Bayesian sample.

To test whether the evolution of malarial parasites was constrained by the spatial distribution of their hosts, we used the software ParaFit (Legendre *et al.*, 2002). The implemented approach follows a matrix exercise that performs a global test of host–parasite cospeciation but can also be used to assess the relative weight of each parasite–host link in mediating congruent evolution. We applied this method to combine the phylogenetic distance matrix of the studied parasite lineages with the geographic distance matrix of their blue tit hosts based on a matrix of incidences of infection (yes or no) between parasites and hosts and then compared this observed matrix with an expected matrix that can be calculated by the randomization of the incidence matrix.

Results

Analyses at the level of sampling: lineage and host effects

We identified 13 different parasite lineages (four belonging to *Parahaemoproteus* and nine belonging to *Plasmodium*, *sensu* Martinsen *et al.*, 2008) across the nine sampling locations. The overall prevalence of avian malaria ranged from 30.5% (Wytham) to 100% (San Ildefonso) (Table 1). We found no difference in overall malaria prevalence between males and females in the sampled host populations (paired *t*-test; $t_8 = -0.389$, $P = 0.707$). The composition of the parasite fauna ranged from one single lineage (San Ildefonso) to seven different lineages (Wytham) (Table 1). For parasite lineages detected in multiple host populations, we found that prevalence was a lineage-specific attribute (two-way ANOVA, effect of lineage: $F_{5,13} = 3.276$, $P = 0.039$) when controlling for differences between sampling sites. This indicates that the same parasite lineage reached a similar prevalence in different host populations. In addition, the model also revealed that the prevalence of parasites was affected by the sampling location (effect of location: $F_{8,13} = 5.906$, $P = 0.003$). Parasite lineage and location explained 84.9% of the variance in parasite prevalence.

Phylogenetic analyses

Phylogenetic constraints on prevalence
When we constrained the evolution of prevalence to be independent of the phylogenetic relationships of lineages by forcing $\lambda = 0$, we found that the H_{mean} of likelihoods converged to a value that was higher than when trait evolution was set to depend strictly on parasite phylogeny by forcing $\lambda = 1$ (H_{mean} after 5 050 000 iterations was 1.639 and −7.809, respectively). The higher H_{mean} value of the first model and the large difference between the two H_{mean} values (BF = 9.448) strongly suggest that a model that assumes phylogenetic independence is better supported than a model that is heavily loaded with phylogenetic inertia (see Methods). This means that prevalence evolved independently of the phylogenetic relationship of the parasite lineages.

Phylogenetic correlations of prevalence and geographical distribution
We investigated how geographical distribution of different lineages as reflected by host population range is related to their prevalence when the phylogenetic history of lineages is taken into account. Accordingly, we tested whether a Bayesian sample of phylogenetic models that allowed a correlation between geographical distribution and prevalence was systematically superior to models that forced the correlation to be zero. The H_{mean} of likelihoods converged to a higher value in the former than in the latter case (H_{mean} after 5 050 000 iterations was −36.525 and −43.320, respectively; this corresponds to BF = 6.795). The posterior distribution of estimated phylogenetic correlations had a mean of 0.799 ± 0.001 (SE). These results imply that across parasite lineages, there is a strong relationship between parasite prevalence and geographical distribution (Fig. 3), which means that more widely distributed parasites infecting several blue tit populations reached higher mean prevalence, whereas

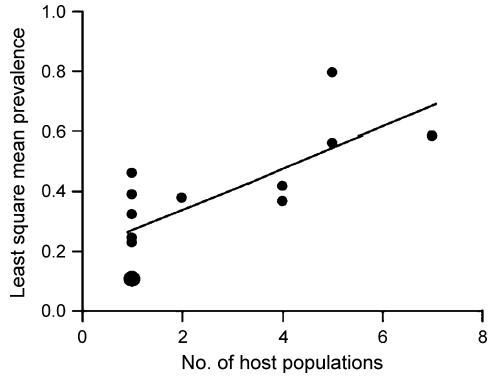

Fig. 3 Least square mean prevalence of malaria lineages in relation to the number of populations in which they were detected. The larger dot indicates two cases.

those found in fewer populations achieved lower prevalence.

We also tested whether distribution range of parasites in terms of the number of blue tit populations in which a parasite was detected correlates with the number of host species that a parasite infects. We found that the posterior distribution of estimated phylogenetic correlations of models that allow correlated trait evolution had a mean of 0.721 ± 0.001 (SE), and such a set of models performs considerably better than models in which the correlation is forced to be zero (BF = 3.525). This means that parasite lineages infecting more blue tit populations have a larger host species range than lineages that were detected in fewer populations (Fig. 4). Using the same phylogenetic approach, we also found a strong correlation between lineage-specific prevalence and host range in terms of the number of host species infected globally (r ± SE = 0.912 ± 0.001, BF = 24.076).

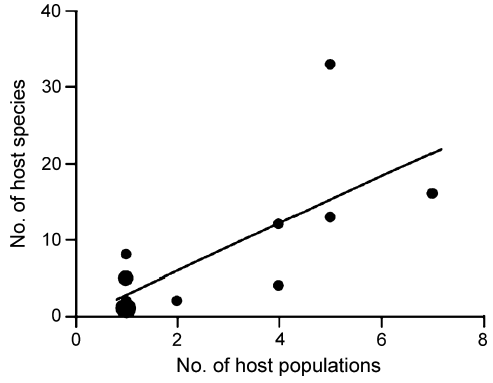

Fig. 4 The number of *Passeriform* host species in relation to the number of blue tit populations in which the parasite lineages were detected. The medium-sized dot indicates two cases, whereas the largest dot indicates three cases.

Spatially constrained evolution

To analyse the historical association between parasite lineages and their host populations, we compared the phylogenetic distances between parasites with the geographic distances between the host populations. The global test of association provided significant evidence for a good match between parasite phylogeny and geographic distances of the host populations (ParaFit global = 3.194, P = 0.008). However, close inspection of individual parasite–host links suggests that not all parasite–host associations contributed equally to the global fit between the two data sets. Of the 34 cases of such association, 13 (38.2%) had probabilities of $P < 0.05$ and one additional link (2.9%) was marginally significant (P = 0.058) (Table 2). Apparently, most parasites

Table 2 Historical association between malaria lineages and blue tit host localities as revealed by the ParaFit results based on naturally occurring infection patterns. The significance of particular host–parasite relationships indicates the degree by which the evolution of the given parasite is constrained by the geographical coordinates of the host populations (see Methods for details). 'Pa' refers to lineages belonging to *Parahaemoproteus*; 'Pl' refers to *Plasmodium* in lineage names. Values in bold indicate statistical significance at $P \leq 0.05$.

Lineage	Location	P
Pl-SW2	Wytham	0.353
Pl-SW2	Revinge	0.564
Pl-SGS1	Antwerp	0.920
Pl-SGS1	Wytham	0.918
Pl-SGS1	**Krakow**	**0.014**
Pl-SGS1	**Vienna**	**0.001**
Pl-SGS1	**Pilisszentlászló**	**0.001**
Pl-GRW11	Wytham	0.915
Pl-GRW11	**Krakow**	**0.016**
Pl-GRW11	**Vienna**	**0.003**
Pl-GRW11	**Pilisszentlászló**	**0.001**
Pl-ACAGR1	**Krakow**	**0.030**
Pl-COLL1	**Vienna**	**0.003**
Pl-BT7	Wytham	0.865
Pl-BT7	Krakow	0.058
Pl-BT7	Hemse	0.521
Pl-BT7	Revinge	0.577
Pl-TURDUS1	Oulu	0.991
Pl-TURDUS1	Wytham	0.868
Pl-TURDUS1	**Krakow**	**0.047**
Pl-TURDUS1	Hemse	0.511
Pl-TURDUS1	Revinge	0.594
Pl-GRW6	**Vienna**	**0.025**
Pl-BLUTI1	Wytham	0.608
Pa-PARUS1	**San Ildefonso**	**0.001**
Pa-PARUS1	**Oulu**	**0.010**
Pa-PARUS1	Antwerp	0.193
Pa-PARUS1	Krakow	0.967
Pa-PARUS1	Vienna	0.999
Pa-PARUS1	Hemse	0.443
Pa-PARUS1	Revinge	0.417
Pa-PARCAE1	**Oulu**	**0.031**
Pa-ROFI1	Revinge	0.587
Pa-BLUTI1	Wytham	0.421

detected in the central European populations (Vienna, Pilisszentlászló, Krakow) were strongly associated with their host populations, whereas those detected in the island populations tended to show weaker associations (Wytham, Hemse) (Table 2).

Discussion

We showed that the prevalence of avian malarial infections differed among blue tit populations. Moreover, prevalence was a lineage-specific attribute and was affected by the distribution of parasite lineages across host populations and host species, so that lineages found in more blue tit populations or in more passerine species were more prevalent within the populations of blue tit. We also showed that parasites that were more widely distributed among blue tit populations were also found to infect more host species. Phylogenetic relationships among the parasites, however, did not predict the prevalence of the lineages. Finally, we found a close match between the geographic distance-based host tree and the phylogenetic tree of the parasites, suggesting that neighbouring host populations shared related parasite fauna.

Variation in habitat characteristics of the sampling sites may affect parasite transmission through the availability of different hosts, vectors and parasite species (see Merino *et al.*, 2008). Hence, our result showing that the prevalence of malaria lineages differed among host populations is not surprising. However, the relationship between geographical distances of host populations and the phylogenetic distances of parasite lineages is more interesting, because marked differences in both the prevalence and the composition of avian malaria fauna have been found even within a population due to small-scale habitat differences (Wood *et al.*, 2007). Therefore, instead of suggesting that habitat variation due to geographical distances is responsible for this relationship, we find it more likely that host–parasite coevolutionary history is shared between neighbouring host populations. Although we have no information about the genetic relatedness of the studied host populations, we assume, based on the previous studies of other species (Nagai *et al.*, 2007; Lindsay *et al.*, 2008) and a study on Belgian blue tit populations (Verheyen *et al.*, 1995), that geographical distances reflect genetic distances between the host populations. If this assumption is true also for the studied blue tit populations, the evolutionary explanation for the observed relationship could be that neighbouring host populations are genetically more similar which may have resulted in similar host–parasite coevolutionary patterns. The underlying genes may involve, for example, the major histocompatibility complex genes that may play an important role in resistance against malaria (Westerdahl *et al.*, 2005; Bonneaud *et al.*, 2006; Bowen *et al.*, 2006).

For malaria lineages detected in multiple blue tit populations, we found that prevalence was a lineage-specific attribute, i.e. the same lineage reached a similar prevalence in different host populations. Notably, this characteristic of the parasites was independent of the phylogenetic relationship between the different lineages. The lack of a phylogenetic signal here may indicate that prevalence is an evolutionary fast-changing attribute or that prevalence evolves independently from phylogeny, which would not be surprising given the coevolutionary arms race between host resistance and parasite virulence. For example, such arms races are thought to have favoured the evolution of the different strains of the human malarial parasite *P. falciparum*, which show antigenic polymorphism helping it to evade the host immune system (Good *et al.*, 1988). However, even if rapid evolution characterizes the evolution of lineages, prevalence varies more conservatively within lineages, which makes prevalence a lineage-specific trait. This is an important requirement to be met for hypotheses that explain detected prevalences based on evolutionary constraints.

Interestingly, only one study has so far investigated the relationship between prevalence and distribution of avian malaria (Hellgren *et al.*, 2009), although this concerned parasite distribution across host species and has not tested for within-lineage variation of prevalence. Despite theoretical considerations suggesting that host generalist parasites may be unable to exploit each host species to the same extent (trade-off hypothesis: Ewald, 1983; Poulin, 1998), because a lineage cannot develop evasive mechanisms against the defences of all hosts, Hellgren *et al.* (2009) showed that avian malaria lineages infecting a wider host range also reached higher prevalence in single host species. However, little is known about whether distribution and adaptation of a parasite species to different populations of the same host species has an effect on parasite prevalence.

Similar to the study by Hellgren *et al.* (2009), we found that the mean prevalence of avian malaria lineages was positively correlated with the number of blue tit populations and the number of bird species in which these parasites were detected. Distribution range of the parasites on the host population and the species level were also correlated with each other, that is parasites detected from more blue tit populations were also found to infect more passerine bird species, providing additional support for the resource breadth hypothesis. One may argue that this correlation is due to the fact that parasites with restricted distribution (both in blue tits and in other host species) cannot simply expand their distribution due to factors limiting their transmission possibilities (e.g. the lack of vectors in other habitats). However, this is not the case, because all parasites that were found in blue tits and that were also reported in the MalAvi database were also detected in migratory birds, but even more importantly 88.9% of these parasites were found all over Europe and also on other continents with different climatic conditions (data from MalAvi). This suggests that conditions

would be suitable for the transmission of parasites, which are specialists in blue tits and also in other habitats. Hence, the distribution of these parasites is probably a consequence of their different host exploitation strategies to maximize their reproductive success. This idea is further supported by the fact that prevalence was a lineage-specific attribute, that is a parasite reached similar prevalence in different populations.

In conclusion, we showed that the distribution and prevalence of avian malarial parasites within a single host species are influenced by multiple factors. The evolution of avian malaria in blue tits was constrained by the geographical distance of their host populations, because there was a strong association between the geographical distance matrix–based host tree and the phylogenetic tree of the parasites. Prevalence was shown to be a lineage-specific attribute and was also influenced by the geographical location of the host population and the geographical distribution of parasites but not by parasite phylogeny. Most importantly, we found that parasites reaching high prevalence in a single population also occurred in more blue tit populations and also infected a broader range of host species. To improve our understanding of host–parasite relationships, further studies are clearly required. Factors causing the observed differences among host populations should be investigated; especially the role of resistance genes in the evolution of parasite prevalence and virulence should be explored. It would also be important to know how distribution and prevalence of parasites are related to virulence and how all these factors affect the life history of hosts.

Acknowledgments

We thank those people who helped in sample collection and analysis, especially SCL Knowles, J Martínez de la Puente, E Lobato, G Hegyi and M Herényi. We also thank H Winkler, R-T Klumpp and A Fojt for logistical support for the Vienna study. The study was supported by a Synthesys grant from the European Union and a grant from the Kungliga Fysiografiska Sällskapet i Lund to ES, grants from the Hungarian Scientific Research Fund (OTKA 75618, F68295, PD75481) to JT and BR, the Swedish Research Council (individual grants to JÅN and DH) and CAnMove (a Linnaeus Research Excellence Centre funded by the Swedish Research Council and Lund University) to JÅN and DH, the Max Planck Society to BK, the Thule Institute in University of Oulu to SR, the University of Antwerp to LZG and ME, the Spanish National Research Council (CSIC, Spain) to LZG, the Spanish Ministry of Science and Innovation (project CGL2009-09439) to SM and LZG, DS/WBiNoZ/INOS/767/10 to MC. MJW was supported by a Natural Environment Research Council standard grant to Prof Ben Sheldon. The European Union and the European Social Fund have provided financial support to the project under the grant agreement no. TÁMOP 4.2.1./B-09/1/KMR-2010-0003. We thank J Donés, Director of Centro Montes de Valsaín, the Dirección General del Medio Natural (Junta de Castilla y León) and the Pilis Park Forestry to authorize the work in the San Ildefonso and Pilisszentlászló study areas.

References

Beadell, J.S., Gering, E., Austin, J., Dumbacher, J.P., Peirce, M.A., Pratt, T.K. *et al.* 2004. Prevalence and differential host-specificity of two avian blood parasite genera in the Australo-Papuan region. *Mol. Ecol.* **13**: 3829–3844.

Beadell, J.S., Covas, R., Gebhard, C., Ishtiaq, F., Melo, M., Schmidt, B.K. *et al.* 2009. Host associations and evolutionary relationships of avian blood parasites from West Africa. *Int. J. Parasitol.* **39**: 257–266.

Bensch, S., Pérez-Tris, J., Waldenström, J. & Hellgren, O. 2004. Linkage between nuclear and mitochondrial DNA sequences in avian malaria parasites-multiple cases of cryptic speciation? *Evolution* **58**: 1617–1621.

Bensch, S., Hellgren, O. & Pérez-Tris, J. 2009. MalAvi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial cytochrome *b* lineages. *Mol. Ecol. Resour.* **9**: 1353–1358.

Bonneaud, C., Pérez-Tris, J., Federici, P., Chastel, O. & Sorci, G. 2006. Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution* **60**: 383–389.

Bowen, L., Aldridge, B.M., DeLong, R., Melin, S., Godinez, C., Zavala, A. *et al.* 2006. MHC gene configuration variation in geographically disparate populations of California sea lions (*Zalophus californianus*). *Mol. Ecol.* **15**: 529–533.

Brown, J.H. 1984. On the relationship between abundance and distribution of species. *Am. Nat.* **124**: 255–279.

Combes, C. 1997. Fitness of parasites: pathology and selection. *Int. J. Parasitol.* **27**: 1–10.

Dybdahl, M.F. & Storfer, A. 2003. Parasite local adaptation: Red Queen versus Suicide King. *Trends Ecol. Evol.* **18**: 523–530.

Escobedo, G., Roberts, C.W., Carrero, J.C. & Morales-Montor, J. 2005. Parasite regulation by host hormones: an old mechanism of host exploitation? *Trends Parasitol.* **21**: 588–593.

Ewald, P.W. 1983. Host–parasite relations, vectors, and the evolution of disease severity. *Annu. Rev. Ecol. Syst.* **14**: 465–485.

Fargallo, J.A. 2004. Latitudinal trends of reproductive traits in the blue tit *Parus caeruleus*. *Ardeola* **51**: 177–190.

Freckleton, R.P., Harvey, P.H. & Pagel, M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **160**: 712–726.

Garamszegi, L.Z. 2006. The evolution of virulence and host specialization in malaria parasites of primates. *Ecol. Lett.* **9**: 933–940.

Garamszegi, L.Z. 2011. The evolution of virulence in primate malaria parasites based on Bayesian reconstructions of ancestral states. *Int. J. Parasitol.* **41**: 205–212.

Good, M.F., Pombo, D., Quakyi, I.A., Riley, E.M., Houghten, R.A., Menon, A. *et al.* 1988. Human T-cell recognition of the circumsporozoite protein of *Plasmodium falciparum*: immunodominant T-cell domains map to the polymorphic regions of the molecule. *Proc. Natl Acad. Sci. USA* **85**: 1199–1203.

Griffiths, R., Double, M.C., Orr, K. & Dawson, R.J.G. 1998. A DNA test to sex most birds. *Mol. Ecol.* **7**: 1071–1075.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95–98.

Hellgren, O., Waldenström, J., Pérez-Tris, J., Szöllősi, E., Hasselquist, D., Krizanauskiene, A. *et al.* 2007. Detecting shifts of transmission areas in avian blood parasites – a phylogenetic approach. *Mol. Ecol.* **16**: 1281–1290.

Hellgren, O., Pérez-Tris, J. & Bensch, S. 2009. A Jack of all trades and still a master of some: prevalence and host range in avian malaria and related blood parasites. *Ecology* **90**: 2840–2849.

Huelsenbeck, J.P. & Ronquist, F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Kimura, M., Dhondt, A.A. & Lovette, I.J. 2006. Phylogeographic structuring of *Plasmodium* lineages across the North American range of the House Finch (*Carpodacus mexicanus*). *J. Parasitol.* **92**: 1043–1049.

Krasnov, B.R., Poulin, R., Shenbrot, G.J., Mouillot, D. & Khokhlova, J.S. 2004. Ectoparasitic ''Jack-of-All-Trades'': relationship between abundance and host specificity in fleas (Siphonaptera) parasitic on small mammals. *Am. Nat.* **164**: 506–516.

Legendre, P., Desdevises, Y. & Bazin, E. 2002. A statistical test for host–parasite coevolution. *Syst. Biol.* **51**: 217–234.

Lindsay, D.L., Barr, K.R., Lance, R.F., Tweddale, S.A., Hayden, T.J. & Leberg, P.L. 2008. Habitat fragmentation and genetic diversity of an endangered, migratory songbird, the golden-cheeked warbler (*Dendroica chrysoparia*). *Mol. Ecol.* **17**: 2122–2133.

Lucas, J.R., Brodin, A., de Kort, S.R. & Clayton, N.S. 2004. Does hippocampal size correlate with the degree of caching specialization? *Proc. Biol. Sci.* **271**: 2423–2429.

Malmqvist, B., Strasevicius, D., Hellgren, O., Adler, P.H. & Bensch, S. 2004. Vertebrate host specificity of wild-caught blackflies revealed by mitochondrial DNA in blood. *Proc. Biol. Sci.* **271**: 152–155.

Martinsen, E.S., Paperna, I. & Schall, J.J. 2006. Morphological versus molecular identification of avian Haemosporidia: an exploration of three species concepts. *Parasitology* **133**: 279–288.

Martinsen, E.S., Perkins, S.L. & Schall, J.J. 2008. A three-genome phylogeny of malaria parasites (*Plasmodium* and closely related genera): evolution of life-history traits and host switches. *Mol. Phylogenet. Evol.* **47**: 261–273.

McCurdy, D.G., Shutler, D., Mullie, A. & Forbes, M. 1998. Sex-biased parasitism of avian hosts: relation to blood parasite taxon and mating system. *Oikos* **82**: 303–312.

Merino, S., Moreno, J., Vásquez, R.A., Martínez, J., Sánchez-Monsálvez, I., Estades, C.F. *et al.* 2008. Haematozoa in forest birds from southern Chile: latitudinal gradients in prevalence and parasite lineage richness. *Austral. Ecol.* **33**: 329–340.

Nagai, S., Lian, C., Yamaguchi, S., Hamaguchi, M., Matsuyama, Y., Itakura, S. *et al.* 2007. Microsatellite markers reveal population genetic structure of the toxic dinoflagellate *Alexandrium tamarense* (*Dinophyceae*). *J. Phycol.* **43**: 43–54.

Nilsson, A.L.K., Alerstam, T. & Nilsson, J.Å. 2008. Diffuse, short and slow migration among Blue Tits. *J. Ornithol.* **149**: 365–373.

Pagel, M. & Meade, A. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am. Nat.* **167**: 808–825.

Pagel, M., Meade, A. & Barker, D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* **53**: 673–684.

Poulin, R. 1998. Large-scale patterns of host use by parasites of fresh-water fishes. *Ecol. Lett.* **1**: 118–128.

Poulin, R. & Mouillot, D. 2004. The relationship between specialization and local abundance: the case of helminth parasites of birds. *Oecologia* **140**: 372–378.

Ricklefs, R.E. & Fallon, S.M. 2002. Diversification and host switching in avian malaria parasites. *Proc. Biol. Sci.* **269**: 885–892.

Ricklefs, R.E., Fallon, S.M. & Bermingham, E. 2004. Evolutionary relationships, cospeciation, and host switching in avian malaria parasites. *Syst. Biol.* **53**: 111–119.

Ronquist, F. & Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.

Seutin, G., White, B.N. & Boag, P.T. 1991. Preservation of avian blood and tissue samples for DNA analyses. *Can. J. Zool.* **69**: 82–90.

Valkiūnas, G. 2005. *Avian Malaria Parasites and Other Haemosporidia*. CRC Press, Boca Raton, FL.

Verheyen, G.R., Kempenaers, B., Adriaensen, F., van den Broeck, M., Matthysen, E., van Broeckhoven, C. *et al.* 1995. The genetic structure of *Parus caeruleus* (blue tit) populations as revealed by minisatellite single locus probes. *Heredity* **75**: 571–577.

Waldenström, J., Bensch, S., Kiboi, S., Hasselquist, D. & Ottosson, U. 2002. Cross-species infection of blood parasites between resident and migratory songbirds in Africa. *Mol. Ecol.* **11**: 1545–1554.

Waldenström, J., Bensch, S., Hasselquist, D. & Östman, Ö. 2004. A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. *J. Parasitol.* **90**: 194–196.

Westerdahl, H., Waldenström, J., Hansson, B., Hasselquist, D., von Schantz, T. & Bensch, S. 2005. Associations between malaria and MHC genes in a migratory songbird. *Proc. Biol. Sci.* **272**: 1511–1518.

Wood, M.J., Cosgrove, C.L., Wilkin, T.A., Knowles, S.C., Day, K.P. & Sheldon, B.C. 2007. Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, *Cyanistes caeruleus*. *J. Anim. Ecol.* **16**: 3263–3273.

Yotoko, K.S.C. & Elisei, C. 2006. Malaria parasites (Apicomplexa, Haematozoea) and their relationship with their hosts: is there an evolutionary cost for the specialization? *J. Zoolog. Syst. Evol. Res.* **44**: 265–273.

Received 24 November 2010; revised 18 May 2011; accepted 20 May 2011